REMARKS

CLAIM REJECTION, 35 USC Paragraph 103

Claims 1-62 were rejected under U.S.C. 103(a) as being unpatentable over *Breese et al.* (U.S. Patent No. 6,006,218) in view of *Hertz et al.* (U.S. Patent No. 5,754,939).

In reply, the Applicants respectfully disagree.

A. GENERAL COMMENTS

What does the present invention teach and claim in independent claims 1 and 32?

The present invention is a method for **predicting user interests** in documents and products using a **learning machine** and **probability measures**. The steps are among others (See claim 1 and 32):

- **transparently monitoring** user interactions;
- using the monitored user actions (note: transparently monitored) for user-specific files;
- **estimate parameters of a learning machine** to define a user model based on user specific files;
- **using the learning machine** (i.e. with user estimated parameters) to **estimate the probability that a document is of interest to a user** (i.e. probability estimates);
- **using the estimated probability** to provide **personalized information** to user.

The Applicants would like to respectfully note that learning can be divided into two parts: (1) **memorization** and (2) **generalization or prediction**.

Ad 1. Memory

Memory refers to what happened in the **past**. A model could be developed that keeps track or score of what happened. For instance, a **user model** could be developed of the scored/tracked items (e.g. which websites **were** visited or which documents **were** looked at). Items could be correlated or similarities could be established (See e.g. *Hertz* Col. 8, line 49; *Hertz* Claim 3).

Using such a model (called knowledge or memory model) one could determine the probability that a user **has seen** or **knows about** an item. Based on this memory, one could determine correlations/similarities/matches (See e.g. *Hertz* Fig. 10 item **1103**; *Hertz* Col. 78 lines 51-52 "... cluster articles based on similarity ... ") with items obtained through a search query. Note such a model is **only applicable** to determine the probability for:

(1) an **individual user**, and

(2) **for that particular item**.

There is no carry over and **no generalization** to other users or other items. Memorization could also be referred to as low-level learning (or limited learning).

More specifically to *Breese*, who teaches that one could determine the **probability that a user knows** about an item (*Breese*: Column 7, lines 1-10, 31-36) – i.e. the user has seen that item in the past. Note knowledge probability (i.e. memory) as in *Breese* IS NOT the same as probability that documents are of interest (i.e. generalization/estimate probability) as in the present application as an artisan would readily appreciate.

In a model one could further make the distinction between application-dependent or application-independent learning. An example of application-dependent learning could be "choose all relevant NY Times articles". An example of application-independent learning could be "choose all relevant NY Times articles **and** find the most important emails, provide personalized search results, etc.". The Applicants assert that *Hertz* teaches the application-dependent approach, whereas the present application is application-independent as defined by elements 1(e) and 1(f) (same for our claim 32).

Classification as an application-**in**dependent approach requires at least two criteria:

(i) "*cross fertilization*" (see present application), i.e. feedback or learning in one application is used to serve all applications. Neither *Hertz* nor *Breese* teach cross-fertilization.

(ii) a user-model can be used for a new personalized application, without the need for application specific learning or initialization. Neither *Hertz* nor *Breese* teach such a generic user model.

To illustrate the application-dependency of *Hertz*, see for instance column 10, lines 10-24 and column 11, lines 3-16. *Hertz* also teaches different sets of attributes for different applications, which makes it obvious that *Hertz* **can't conceive** an application-independent user model. It is again further noted that the present application **does not teach memorization**. Rather, the present invention teaches a learning model to estimate probabilities to predict personalized information that is of interest to the user.

Ad 2.Generalization

Neither *Breese* nor *Hertz* teach any type of generalization; there is no learning involved other than keeping score or tracking what happened in the past. Please note that there is no learning or generalization in these prior art references and could therefore not suggest the present invention to render it obvious.

For example could *Breese* or *Hertz* use a user-model for apples to predict if the user is interested in pears? The answer is no, since the user-model for apples has no knowledge or generalization power related to pears. The teachings of *Breese* and *Hertz* are knowledge-based without any teaching on how to use that knowledge model to generalize beyond that or become application independent – independent from the apples and extend to pears. It is one of the objectives of the present invention to overcome this shortcoming; i.e. a **learning machine in the probability domain** and **cross-fertilization of learning in one mode to another mode**.

Generalization **predicts beyond** items in the past and even beyond the user itself; it **estimates probability** of something to happen in the future. It is exactly this generalization that is claimed in claims 1 and 32 by:

(1) using the monitored actions to **estimate parameters of a learning machine**, and

(2) using the learning machine **to estimate the probability** that a document is of interest to a user.

As clearly taught in the present application, **generalization** is made possible by defining a model in the **probability domain**, which **decouples particular feature vectors** and learns to make the model **application/item independent**. The user model of the learning machine in the present invention represents user interests **independent** of any specific (note: specific is application dependent) user information. In other words, the present invention is not related to a specific query. There is therefore no need to distinguish between seen or unseen documents.

Furthermore, *Hertz* (Col. 5, lines 4-21) teaches ordering articles. The question arises what the importance is of the ordered articles. For instance, is it important enough to drag your boss out of a meeting to show the article? *Hertz* does not have a solution for this problem. Ordering articles could be useless if on one day the article is of high importance and the next day is of low importance. This is in contrast to the present invention, which determines for every document an absolute score of importance, e.g. 0.9 probability that a document is of interest to a user, independent what the other documents on today's list were. This aspect is clearly claimed in element 1(e) and 1(f) (vice versa in claim 32) of the present application.

Accordingly, the Applicants submit that the present claims 1-62 are **NOT obvious** with respect to *Breese* in view of *Hertz*. A ***prima facie*** case of obviousness (See MPEP 2143) has **not been established** as discussed *supra*.

B. SPECIFIC COMMENTS

Claims 1 and 32

1. The Office Action asserts that column 5, lines 25-38 of *Breese* discloses, "transparently monitoring user interactions with data while the user is engaged in normal use of a computer."

In reply, the Applicants assert that the cited passages **do not** specify **nor imply** that the user is engaged in normal use of the computer, nor that the monitoring is transparent. In fact, the cited passage includes obtaining information from questionnaire results, which are certainly not transparently obtained when the user is engaged in normal use of a computer.

2. The Office Action asserts that column 8, lines 33-36, 44-46 of *Breese* discloses, "updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user."

In reply, the Applicants assert that if the step in element (a) "transparently monitoring user interactions .." is not taught or implied, then there can not be a teaching or implication of step (b) that follows (a). Note it is updating (step b) with the monitored user interactions (step a).

3. The Office Action asserts that element, "analyzing a document to identify properties of the document," is described in column 8, lines 15-26 of *Breese*.

In reply, the cited section of *Breese* does not discuss any analysis of documents and are irrelevant to the claim element.

4. The Office Action asserts that several sections of *Hertz* discloses steps (c), (e) and (f).

In reply, the Applicants respectfully disagree and refer to the arguments made *supra* (general comments). The Applicants would like to respectfully point out that the Office Action fails to **clearly point out** where *Hertz* teaches steps (c), (e) and (f) since reviewing these sections the Applications are unable to identify the relevant teachings. Perhaps the Examiner could assist and be more precise by pointing to the selective sentences instead of an *aggregate* of independent sections/paragraphs/words.

In addition, *Hertz*:

(i) teaches **memorization**, we don't,

(ii) teaches an **application specific** user model without any generalization power, we have an application-independent learning model,

(iii) does not teach or imply any learning to **estimate probability** of user interests, we do,

(iv) does not teach or imply any information theory to **determine probability measures**, we do,

(v) does not teach **probability measures** if whether an item is of interest to a user (See also *infra*), we do, and/or

(vi) teaches **clusters of documents** (See Hertz Col. 78, lines 51-53) and does not teach **clusters of user models** like we do (which is a big difference).

None of the sections (either *individually* or *combined*) of *Hertz* referred to in the Office Action discusses, teaches or implies steps (either *individually* or *combined*) (c), (e) and (f). Accordingly, the Applicants submit, as submitted *supra*, that the present claims 1-62 are **NOT obvious** with respect to *Breese* in view of *Hertz*. A ***prima facie*** case of obviousness (See MPEP 2143) has **not been established**.

CLAIMS 2-31 and 33-62

The Applicants believe that the significant differences discussed above between the claimed invention and *Breese* in view of *Hertz* make the claimed invention novel and *non*-obvious. Because all other claims depend from either claim 1 or claim 32, the Applicants believe that all pending depending claims are also novel and *non*-obvious. In addition to their dependency on claims 1 or 32, the Applicants incorporate herewith **all previous arguments** made on the record in the previous reply to the first Office Action.

In addition, the Applicants have trouble comprehending the relevant teaches pointed out by the Examiner related to *Hertz* that would render the present claims obvious. As a side note, *Hertz* in Column 7, lines 47-67 to Column 8 1-9 teaches "truly passive" and "browsing and filtering", which shows that *Hertz* does not have the intention to suggest its teachings to be a basis for predicting user interests for personal search and services. This is in contrast to claim 1 and 32 of the present application.

Furthermore, Applicants would like to point out that *Hertz* does not teach nor imply probability measures, or how to define probability measures in either formula or

wordings. A simple word search on the word probability in *Hertz* doesn't return a favorable answer. Note the word "probability" can be found e.g. in *Hertz* Col. 50 line 28 it refers to " ... probability that a user will access target object T". **However**, this probability is based on a memorized user model (see *supra*) and not the probability that the document is of interest to a user (which is based on a learning model of estimated probabilities and not memories). Furthermore, a description or implication of the necessary information theory to establish probability measures as claimed in claim 1 and 32 is missing in *Hertz*. Accordingly, the Applicants are puzzled to why the Office Action asserts that *Hertz* teaches or renders our claims obvious in combination with *Breese*.

CONCLUSION

Applicants respectfully submit that the present claims 1-62 are **NOT obvious** with respect to *Breese* in view of *Hertz*. A ***prima facie*** case of obviousness (MPEP 2143) has **not been established** as discussed *supra*. Even *if* at the time the invention (i.e. hindsight is impermissible, See MPEP 2141.01 III) was made one skilled in the art would be motivated to combine *Breese* and *Hertz*, the resulting method would still not possess the capability to provide automated and personalized information services to a user that uses machine learning including memorization and generalization defined in the probability domain simply because neither *Breese* or *Hertz* teach or suggest anything beyond memorization models.

Therefore, the Applicants submit that claims 1-62 are novel and unobvious over the closest prior art of record. Accordingly, allowance of the claims now in the application is kindly requested.

Respectfully submitted,

Dr. Ron Jacobs
Reg. No. 50,142
LUMEN Intellectual Property Services
2345 Yale Street, 2nd Floor
Palo Alto, CA 94306-1429

Phone: (650) 424-0100
Fax: (650) 424-0141
Email: ron@lumen.com